Facial Reduction and Partial Polyhedrality

Bruno F. Lourenço ∗ Masakazu Muramatsu† Takashi Tsuchiya ‡

December 2015

Abstract

In this article we present FRA-Poly, a facial reduction algorithm (FRA) for conic linear programs that is sensitive to the presence of polyhedral faces in the cone. The main goals of FRA and FRA-Poly are the same, i.e., finding the minimal face containing the feasible region and detecting infeasibility, but FRA-Poly treats polyhedral constraints separately. This idea enables us to reduce the number of iterations drastically when there are many linear inequality constraints. The worst case number of iterations for FRA-poly is written in the terms of a “distance to polyhedrality” quantity and provides better bounds than FRA under mild conditions. In particular, in the case of the doubly nonnegative cone, FRA-Poly gives a worst case bound of \( n \) whereas the classical FRA is \( O(n^2) \). Of possible independent interest, we also prove a variant of Gordan-Stiemke’s Theorem and of a proper separation theorem that takes into account the presence of partial polyhedrality. We also present a few applications of our approach. In particular, we will use FRA-poly to improve the bounds recently obtained by Liu and Pataki on the dimension of certain sub-affine spaces which appear in weakly infeasible problems.

1 Introduction

Consider the following pair of primal and dual optimization problems:

\[
\begin{align*}
\inf_x \ & \langle c, x \rangle \\
\text{subject to} \ & A x = b \\
& x \in \mathcal{K}^* \\
\sup_y \ & \langle b, y \rangle \\
\text{subject to} \ & c - A^T y \in \mathcal{K},
\end{align*}
\]

where \( \mathcal{K} \subseteq \mathbb{R}^n \) is a closed convex cone and \( \mathcal{K}^* \) is the dual cone \( \{ s \in \mathbb{R}^n \mid \langle s, x \rangle \geq 0, \forall x \in \mathcal{K} \} \). We have that \( A : \mathbb{R}^n \to \mathbb{R}^m \) is a linear map, \( b \in \mathbb{R}^m \), \( c \in \mathbb{R}^n \) and \( A^T \) denotes the adjoint map. We also have \( A^T y = \sum_{i=1}^m A_i y_i \), for certain elements \( A_i \in \mathbb{R}^n \). The inner product is denoted by \( \langle \cdot , \cdot \rangle \). We will use \( \theta_P \) and \( \theta_D \) to denote the primal and dual optimal value, respectively.

In the absence of either a primal relative interior feasible solution or a dual relative interior slack, it is possible that \( \theta_P \neq \theta_D \). A possible way of correcting that is to let \( \mathcal{F}_{\text{min}}^D \) be the minimal face of \( \mathcal{K} \) which

\[\text{Minimize } \langle c, x \rangle \text{ subject to } A x = b, x \in \mathcal{K}^* \]
contains the feasible slacks \( F_D^i = \{ c - A^Ty \in K \mid y \in \mathbb{R}^m \} \), then we substitute \( K \) by \( F_{\min}^D \) and \( K^\ast \) by \( (F_{\min}^D)^\ast \). With that, the new primal optimal value \( \theta_{p'} \) will satisfy \( \theta_{p'} = \theta_D \). This is precisely what facial reduction [5, 19, 25] approaches do.

In this paper, we analyze how to take advantage of the presence of polyhedral faces in \( K \) when doing Facial Reduction. To do that, we introduce FRA-Poly, which is a facial reduction algorithm (FRA) that, in many cases, provides a better worst case complexity than the usual approach, specially when \( K \) is a direct product of several cones. The idea behind it is as follows. Facial reduction algorithms work by successively identifying what is called “reducing directions” \( \{d_1, \ldots, d_t\} \). Starting with \( F_1 = K \), these directions define faces of \( K \) by the relation \( F_{i+1} = F_i \cap \{d_i\}^\perp \). Of course, these directions are not arbitrary and for feasible problems, \( d_i \) must be such that \( F_{i+1} \) is a face of \( K \) containing \( F_D^i \). We then obtain a sequence \( F_1 \supseteq \ldots \supseteq F_t \) of faces of \( K \) such that \( F_i \supseteq F_D^i \) for every \( i \). Usually, a FRA proceeds until \( F_{\min}^D \) is found.

A key observation is that as soon as we reach a polyhedral face \( F_i \), we can jump to the minimal face \( F_{\min}^D \) in a single facial reduction step. In addition, when \( K \) is a direct product \( K = K^1 \times \ldots \times K^r \), each \( F_i \) is also a direct product \( F_i^1 \times \ldots \times F_i^r \). In this case an even weaker condition is sufficient to jump to \( F_{\min}^D \), namely, if every block \( F_i^j \) satisfy one of the two conditions: a) \( F_i^j \) is polyhedral; b) every \( s \in F_i^j \cap \ker A \cap \{c\}^\perp \) is such that \( s^j \in (F_i^j)^\perp \). When the problem is feasible, the condition b) is equivalent to the statement that the \( j \)-th block of \( F_{\min}^D \) is equal to \( F_i^j \). In summary, we can jump to the minimal face if either a block is polyhedral or it is already a piece of the minimal face.

Our proposed algorithm FRA-Poly works in two phases. In Phase 1, it proceeds until a face \( F_i \) satisfying the condition above is reached or until a certificate of infeasibility is found. To do that, we model the search for reducing directions as a pair of primal and dual problems satisfying a generalized Slater’s condition that takes into account polyhedrality. In Phase 2, \( F_{\min}^D \) is obtained with single facial reduction step. One interesting point is that even if \( F_i \) is different from the minimal face, it is possible to show that if we reformulate [D] as a problem over \( F_i \), then strong duality will hold. The theoretical backing for that is given in Section 4 where we discuss a classical strong duality criterion together with a generalization of the Gordan-Stiemke Theorem for the case when \( K \) is the direct product a closed convex cone and a polyhedral cone, see Theorem 5. We also prove a proper separation theorem that will be the engine behind FRA-poly, see Theorem 8.

In order to analyze the number of facial reduction steps in the worst case, we introduce a quantity called \textit{distance to polyhedrality} \( \ell(\mathbb{K}) \). This is the length minus one of the longest strictly ascending chain of nonempty faces \( F_i \supseteq \ldots \supseteq F_t \) for which \( F_1 \) is polyhedral and \( F_t \) is not polyhedral for all \( i > 1 \). If \( K \) is a direct product of arbitrary cones \( K^1 \times \ldots \times K^r \), we prove that FRA-Poly stops in at most \( 1 + \sum_{i=1}^r \ell(\mathbb{K}^i) \) steps. This quantity is no worse than the bound given by classical FRA and, provided that at least two of the cones \( K^i \) are not subspaces, it is strictly smaller.

As an application, we give a nontrivial bound for the singularity degree of CLPs over cones that are intersections of two other cones, such as the doubly nonnegative cone \( D^n \). For the case \( K = D^n \), we show that the longest chain of nonempty faces of \( D^n \) has length \( 1 + \frac{n(n+1)}{2} \), which is the maximum allowed for a cone contained in the space \( S^n \) of \( n \times n \) symmetric matrices. Therefore, the classical analysis gives the upper bound \( \frac{n(n+1)}{2} \) for the singularity degree of feasible problems over the doubly nonnegative cone. On the other hand, using our technique, we show that the singularity degree of any problem over \( D^n \) is at most \( n \). We also use FRA-poly to improve the bounds obtained by Liu and Pataki in Theorem 9 of [10] on the dimension of certain subspaces connected to weakly infeasible problems.

Table 1 contains a summary of the bounds predicted by FRA and FRA-poly for several cases. The notation \( \ell_K \) indicates the length of the longest strictly ascending chain of nonempty faces of \( K \). The first line correspond to a single cone, the second to a product of \( r \) arbitrary closed convex cones and the third to the product of \( r_1 \) Lorentz (quadratic) cones and \( r_2 \) positive semidefinite cones, respectively. These results follow from Proposition 1 and Example 14. The last line contains the bounds for the doubly nonnegative cone, which follows from Proposition 26 and Corollary 25.

This work is divided as follows. In Section 2 we give some background on related notions including a brief discussion on strict complementarity. In Section 3 we review facial reduction and discuss a pair of auxiliary problems that will be used in FRA-Poly. In Section 4 we prove versions of two classical theorems
the elements of \( F \) space" program (D) can be in four different feasibility statuses: strongly infeasible if \( \text{dist}(c, F) = \infty \), feasible if \( \text{dist}(c, F) = 0 \), and weakly feasible if \( \text{dist}(c, F) > 0 \), in which case the dual optimal value is \( 0 \). No duality gap is present.

We recall two important notions of separation. Two convex sets can be properly separated if and only if they are not both contained in \( H \) but they are not both contained in \( H \). Two convex sets can be properly separated if and only if their relative interiors are disjoint.

Let \( \mathcal{K} \) be a closed convex cone, we write \( \mathcal{K}^* \) for its dual cone. We write \( \text{span} \mathcal{K} \) for its linear span and \( \text{lin} \mathcal{K} \) for its lineality space, which is \( \mathcal{K} \cap -\mathcal{K} \). If \( \mathcal{F} \) is a face of \( \mathcal{K} \), we define the conjugated face of \( \mathcal{F} \) as \( \mathcal{F}^\Delta = \mathcal{K}^* \cap F^\perp \). If we select a point in the relative interior \( x \) of \( \mathcal{F} \), we have \( \mathcal{F}^\Delta = \mathcal{K}^* \cap \{ x \}^\perp \). Note that the elements of \( \mathcal{F}^\Delta \) define the supporting hyperplanes of \( \mathcal{K} \) passing through \( x \). Given \( z \in \mathcal{K} \), we will use the notation \( \mathcal{F}(z, \mathcal{K}) \) to denote the minimal face of \( \mathcal{K} \) which contains \( z \).

We denote the dual feasible region by \( \mathcal{F}_D = \{ y \in \mathbb{R}^n \mid c - A^T y \in \mathcal{K} \} \). We also write \( \mathcal{F}_D^\Delta \) for the "slack space" \( \mathcal{F}_D^\Delta = \{ c - A^T y \mid y \in \mathcal{F}_D \} \). The primal feasible region is \( \mathcal{F}_P = \{ x \in \mathcal{K}^* \mid Ax = b \} \). The conic linear program (D) can be in four different feasibility statuses: 

1. Strongly feasible if \( \mathcal{F}_D^\Delta \cap \text{ri} \mathcal{K} = \emptyset \); 
2. Weakly feasible if \( \mathcal{F}_D^\Delta \cap \text{ri} \mathcal{K} \neq \emptyset \) but \( \mathcal{F}_P \cap \text{ri} \mathcal{K} = \emptyset \); 
3. Weakly infeasible if \( \mathcal{F}_D^\Delta = \emptyset \) but \( \text{dist}(c + 
abla^2 \mathcal{K}, \mathcal{K}) = 0 \); 
4. Strongly infeasible if \( \text{dist}(c + \nabla^2 \mathcal{K}, \mathcal{K}) > 0 \).

Denote the dual optimal value by \( \theta_D \) and the primal optimal value by \( \theta_P \). The strong duality theorem states that if \( \mathcal{F}_D^\Delta \) is strongly feasible and \( \theta_D < +\infty \), then \( \theta_P = \theta_D \) and \( \theta_P = \theta_D \) is attained. On the other hand, if \( \mathcal{F}_P \) is strongly feasible and \( \theta_P > -\infty \), then \( \theta_P = \theta_D \) and \( \theta_P = \theta_D \) is attained.

Suppose that \( x \) and \( y \) are optimal solutions for (P) and (D) and that there is no duality gap. Let \( s \) be the corresponding optimal slack \( c - A^T y \). Then the equation \( (x, s) = 0 \) holds. This implies that \( x \in \mathcal{F}(s, \mathcal{K})^\Delta \) and \( s \in \mathcal{F}(x, \mathcal{K}^*)^\Delta \). We have then competing definitions of strict complementarity. The next proposition shows the relation between them.

Proposition 1. Let \( s \in \mathcal{K} \) and \( x \in \mathcal{K}^* \). Consider the following statements:

1. \( \mathcal{F}(s, \mathcal{K})^\Delta = \mathcal{F}(x, \mathcal{K}^*) \);
2. \( \mathcal{F}(s, \mathcal{K}) = \mathcal{F}(x, \mathcal{K}^*)^\Delta \).

(Pataki [13], see Remark 3.6 therein).
ii. (a) there exists a face $\mathcal{F} \subseteq \mathcal{K}$ such that $s \in \text{ri} \mathcal{F}$ and $x \in \text{ri} \mathcal{F}^\Delta$;
(b) there exists face $\mathcal{F} \subseteq \mathcal{K}^*$ such that $x \in \text{ri} \mathcal{F}$ and $s \in \text{ri} \mathcal{F}^\Delta$.

(see Section 2 in Tunçel and Wolkowicz [24]);

(iii. (a) $x \in \text{ri} \mathcal{F}(s, \mathcal{K})^\Delta$;
(b) $s \in \text{ri} \mathcal{F}(x, \mathcal{K}^*)^\Delta$.

Items i.a), ii.a), iii.a) are equivalent. Items i.b), ii.b), iii.b) are also equivalent. If $\mathcal{K}$ and $\mathcal{K}^*$ are facially exposed, then iii.a) and iii.b) are also equivalent.

Proof. We will only prove one set of equivalences since the other has a similar proof.

(i.a) $\Rightarrow$ ii.a) Suppose that $\mathcal{F}(s, \mathcal{K})^\Delta = \mathcal{F}(x, \mathcal{K}^*)$ holds. Take $\mathcal{F} = \mathcal{F}(s, \mathcal{K})$ and recall that the minimal face which contains a point $x$ is characterized by the fact that $x$ belongs to the relative interior of that face. Then, we must have $x \in \text{ri} \mathcal{F}(x, \mathcal{K}^*)$, which implies $x \in \text{ri} \mathcal{F}^\Delta$.

(ii.a) $\Rightarrow$ iii.a) We have $\mathcal{F}(s, \mathcal{K}) = \mathcal{F}$, so that $x \in \text{ri} \mathcal{F}(s, \mathcal{K})^\Delta$.

(iii.a) $\Rightarrow$ i.a) Suppose $x \in \text{ri} \mathcal{F}(s, \mathcal{K})^\Delta$. Because $\mathcal{F}(s, \mathcal{K})^\Delta$ is a face of $\mathcal{K}^*$, we have $\mathcal{F}(s, \mathcal{K})^\Delta = \mathcal{F}(x, \mathcal{K}^*)$.

(iii.a) $\iff$ iii.b) Suppose that both $\mathcal{K}$ and $\mathcal{K}^*$ are facially exposed. Then $\mathcal{F}(s, \mathcal{K})^\Delta = \mathcal{F}(x, \mathcal{K}^*)$ implies $\mathcal{F}(s, \mathcal{K})^{\Delta \Delta} = \mathcal{F}(x, \mathcal{K}^*)^{\Delta \Delta}$. Since $\mathcal{K}$ is facially exposed, we have $\mathcal{F}(s, \mathcal{K}) = \mathcal{F}(s, \mathcal{K})^{\Delta \Delta}$. If the second alternative holds, we have that facial exposedness of $\mathcal{K}^*$ implies $\mathcal{F}(x, \mathcal{K}^*)^{\Delta \Delta} = \mathcal{F}(x, \mathcal{K}^*) = \mathcal{F}(s, \mathcal{K})$.

If items i.a), ii.a), iii.a) hold, then the pair $(x, s)$ is said to be primal strict complementarity. If i.b), ii.b), iii.b) hold, then we have dual strict complementarity. This distinction only matters when $\mathcal{K}$ or $\mathcal{K}^*$ is not facially exposed. We remark that when $\mathcal{K} = \mathbb{R}_+^n$, the notion of strict complementarity above is equivalent to $x + s > 0$. When $\mathcal{K} = \mathbb{S}_+^n$, it is equivalent to $x + s$ being positive definite. More generally, if $\mathcal{K}$ is a symmetric cone, then strict complementarity is equivalent to $x + s \in \text{ri} \mathcal{K}$.

3 Facial Reduction

Facial Reduction was developed by Borwein and Wolkowicz to restore strong duality in convex optimization [4, 5]. Descriptions for the conic linear programming case have appeared, for instance, in Pataki [19] and in Waki and Muramatsu [22].

Here, we will suppose that our main interest is in the dual problem [D]. Let $\mathcal{L} = \text{range } \mathcal{A}^T$. Facial Reduction hinges on the fact that $(\text{ri} \mathcal{K}) \cap (\mathcal{L} + c) = \emptyset$ if and only if there is $d \in \mathcal{K}^*$ such that $\mathcal{A}d = 0$ (i.e., $d \in \mathcal{L}^\perp$) and one of the two alternatives holds: i. $\langle c, d \rangle = 0$ and $d \notin \mathcal{K}^\perp$; or ii. $\langle d, c \rangle < 0$. If such a $d$ exists and $\langle d, c \rangle = 0$ then we have that $\mathcal{F} = \mathcal{K} \cap \{d\}^\perp$ is a proper face of $\mathcal{K}$ which contains the dual feasible region $\mathcal{F}_D^\ast$. If $\langle d, c \rangle < 0$, then the problem must be infeasible. We then substitute $\mathcal{K}$ by $\mathcal{F}$ and repeat. We write below a generic facial reduction algorithm similar to the one described in [25].

[Generic Facial Reduction]

**Input:** $[D]$

**Output:** A set of reducing directions $\{d_1, \ldots, d_t\}$ and $\mathcal{F}_\text{min}^D$.

1. $\mathcal{F}_1 \leftarrow \mathcal{K}, i \leftarrow 1$
2. Let $d_i$ be an element in $\mathcal{F}_i^\perp \cap \ker \mathcal{A}$ such that either: i) $d_i \notin \mathcal{F}_i^\perp$ and $\langle c, x \rangle = 0$; or ii) $\langle c, x \rangle < 0$. If no such $d_i$ exists, let $\mathcal{F}_\text{min}^D \leftarrow \mathcal{F}_i$ and stop.
3. If $\langle c, d_i \rangle < 0$, let $\mathcal{F}_\text{min}^D \leftarrow \emptyset$ and stop.
4. If $\langle c, d_i \rangle = 0$, let $\mathcal{F}_{i+1} \leftarrow \mathcal{F}_i \cap \{d_i\}^\perp, i \leftarrow i + 1$ and return to 2).
In this article, we will refer to the directions satisfying the conditions in Step 2 as reducing directions. An important issue when doing facial reduction is how to model the search for the reducing directions. It is sometimes said that doing facial reduction can be as hard as solving the original problem. However, an important difference is that the search for the reducing directions can be cast as a pair of primal and dual problems which are always strongly feasible. This was shown in the work by Cheung, Schurr and Wolkowicz [6] and in our previous work [12]. Recently, it was shown by Permenter, Friberg and Andersen that reducing directions can also be obtained as by-products of self-dual homogeneous methods, see [20] for details on this connection. There are also approximate approaches such as the one described by Permenter and Parrilo [21], where the search for the reducing directions is conducted in a more tractable cone at the cost of, perhaps, failing to identify $F_{\text{Dmin}}$, but still simplifying the problem nonetheless. In this work, we will make extensive use of the following pair of primal and dual problems.

$$\begin{align*}
\text{minimize} & \quad t \\
\text{subject to} & \quad \langle c, x - te^* + t - w \rangle = 0 \quad (P_K) \\
& \quad \langle e, x \rangle + w = 1 \\
& \quad Ax - tA e^* = 0 \\
& \quad (x, t, w) \in K^* \times \mathbb{R}_+ \times \mathbb{R}_+ \\
\text{maximize} & \quad y_2 \\
\text{subject to} & \quad cy_1 - ey_2 - A^T y_3 \in K \quad (D_K) \\
& \quad 1 - y_1(1 + \langle c, e^* \rangle) + \langle e^*, A^T y_3 \rangle \geq 0 \\
& \quad y_1 - y_2 \geq 0
\end{align*}$$

The pair $(P_K)$ and $(D_K)$ are parametrized by $A$, $c$, $K$ and by fixed elements $e \in K$ and $e^* \in K^*$. We have the following lemma, which suggests a good choice of parameters $e$ and $e^*$, although we will make a different choice of $e$ later.

**Lemma 2.** Consider the pair $(P_K)$ and $(D_K)$ with $e$ and $e^*$ such that $e \in \text{ri} K$ and $e^* \in \text{ri} K^*$. The following properties hold.

i. Both $(P_K)$ and $(D_K)$ have relative interior feasible points.

ii. Let $(x^*, t^*, w^*)$ be a primal optimal solution. The optimal value is zero if and only if $F_{\text{Dmin}} \subseteq K$. Moreover, if the optimal value is zero, we have $\langle c, x^* \rangle < 0$ and $F_{\text{D}} = F_{\text{Dmin}} = \emptyset$, or $\langle c, x^* \rangle = 0$ and $F_{\text{D}} \subseteq K \cap \{x^*\}^\perp \subseteq K$.

iii. Let $(y_1^*, y_2^*, y_3^*)$ be a dual optimal solution. If the common optimal value is nonzero, then $F_{\text{Dmin}} = K$ and $s = c - A^T y^*_3 y^*_1$ is a dual optimal solution satisfying $s \in (\text{ri} F_{\text{D}}^* ) \cap \text{ri} K$.

**Proof.** It is a special case of Lemma 7. See also Lemma 10 in [12].

## 4 Partial Polyhedrality Theorems

In Rockafellar’s classic book [22], among its many notable chapters there is one on “Applications of Polyhedral Convexity”. Many classical results are proved under weaker conditions if some of the objects involved are polyhedral. Borwein and Lewis also include a discussion of the mixed Fenchel Duality Theorem in Section 5.1 of [3]. As far as we could dig in the literature, there does not seem to be many papers that deal with
this subject apart from the ones by Klee such as [9], which is referenced in Rockafellar’s book. The results in this section have a similar flavor.

Before we proceed, we need the following special version of Slater’s condition which takes into account partial polyhedrality. As we could not find a precise reference for it, we give a proof in the Appendix A. It can also be proved by invoking in an appropriate manner a version of Fenchel’s duality theorem that takes into account polyhedrality, such as Theorem 31.1 in [22] or Corollary 5.1.9 in [3]. First, we need the following definition.

**Definition 3** (Partial Polyhedral Slater’s condition). Let \( K = K^1 \times K^2 \), where \( K^1 \subseteq \mathbb{R}^{n_1}, K^2 \subseteq \mathbb{R}^{n_2} \) are closed convex cones such that \( K^2 \) is polyhedral. We say that \( [\mathcal{D}] \) satisfies the Partial Polyhedral Slater’s (PPS) condition if there is a slack \((s_1, s_2) = c - A^T y\), such that \( s_1 \in \text{ri} K^1 \) and \( s_2 \in K^2 \). Similarly, we say that \( [\mathcal{D}] \) satisfies the PPS condition, if there is a primal feasible solution \( x = (x_1, x_2) \) for which \( x_1 \in \text{ri} (K^1)^* \).

**Proposition 4.** Let \( K = K^1 \times K^2 \), where \( K^1 \subseteq \mathbb{R}^{n_1}, K^2 \subseteq \mathbb{R}^{n_2} \) are closed convex cones such that \( K^2 \) is polyhedral.

i) If \( \theta_P \) is finite and \( [\mathcal{D}] \) satisfies the PPS condition, then \( \theta_P = \theta_D \) and the dual optimal value is attained.

ii) If \( \theta_D \) is finite and \( [\mathcal{D}] \) satisfies the PPS condition, then \( \theta_P = \theta_D \) and the primal optimal value is attained.

We now prove a version of the Gordan-Stiemke’s Theorem that takes into account partial polyhedrality. To the best of our knowledge, it is a new result.

**Theorem 5** (Partial Polyhedral Gordan-Stiemke’s Theorem). Let \( \mathcal{L} \) be a subspace and \( K = K^1 \times K^2 \) be a closed convex cone, such that \( K^2 \) is polyhedral. Then we have:

\[
\mathcal{L} \cap K \subseteq (\text{lin} K^1) \times K^2 \iff \mathcal{L}^\perp \cap \left( (\text{ri} K^1)^* \times (K^2)^* \right) \neq \emptyset.
\]

**Proof.** The “\( \Rightarrow \)" implication is straightforward as follows. If \( s = (s_1, s_2) \) belongs to \( \mathcal{L}^\perp \cap \left( (\text{ri} K^1)^* \times (K^2)^* \right) \) and \( x = (x_1, x_2) \) to \( \mathcal{L} \cap K \), then we must have \( \langle x_1, s_1 \rangle = 0 \), which forces \( x_1 \in \text{lin} K^1 \), since \( s_1 \) is a relative interior point.

Next, we prove the “\( \Leftarrow \)" implication. Select a linear map \( A \) such that \( \mathcal{L} = \ker A \). Let \( e^* \in (\text{ri} (K^1)^*) \times \{0\} \) and \( e \in \text{ri} K \). Now consider the following pair of primal and dual problems:

\[
\begin{align*}
\text{minimize} & \quad t && \\
\text{subject to} & \quad \langle e^*, x \rangle + t = 1 \\
& \quad A x - t e = 0 \\
& \quad (x, t) \in K \times \mathbb{R}_+ \\
\end{align*}
\]

\[
\begin{align*}
\text{maximize} & \quad y_1 \\
\text{subject to} & \quad - e^* y_1 - A^T y_2 \in K^* \\
& \quad 1 - y_1 + \langle e, A^T y_2 \rangle \geq 0
\end{align*}
\]

Taking \( \left( \frac{e^*}{\langle e^*, e \rangle}, \frac{1}{\langle e^*, e \rangle} \right) \), we see that the primal problem \( \text{(P}_{GS} \text{)} \) has a relative interior feasible solution. The dual problem \( \text{(D}_{GS} \text{)} \) satisfies the PPS condition and to see that, it is enough to take \( y_1 = -1 \) and \( y_2 = 0 \). This means that both \( \text{(P}_{GS} \text{)} \) and \( \text{(D}_{GS} \text{)} \) attain the optimal value and duality gap is zero. Let \( (x^*, t^*) \) be an optimal solution for \( \text{(P}_{GS} \text{)} \). If \( t^* = 0 \), then \( A x^* = 0, x^* \in K \) and \( \langle e^*, x^* \rangle = 1 \). However, due to our hypothesis, \( x^* \in (\text{lin} K^1) \times K^2 \), which implies that \( \langle e^*, x^* \rangle = 0 \). This is a contradiction, so we must have \( t^* > 0 \) instead. Since \( \text{(D}_{GS} \text{)} \) is attained as well, we have an optimal solution \( (y_1^*, y_2^*) \), with \( y_1^* > 0 \). Because \( - e^* y_1 - A^T y_2 \in K^* \), it readily follows that \( - A^T y_2 \in (\text{ri} (K^1)^*) \times K^2^* \).

\( \square \)
For comparison, we state the classical Gordan-Stiemke’s theorem. Its proof follows from Theorem 5. See also Corollary 2 in Luo, Sturm and Zhang [1].

**Theorem 6** (Gordan-Stiemke’s Theorem). Let \( \mathcal{L} \) be a subspace and \( \mathcal{K} \) be a closed convex cone. Then we have:
\[
\mathcal{L} \cap \mathcal{K} \subseteq \text{lin } \mathcal{K} \iff \mathcal{L}^\perp \cap (\text{ri } \mathcal{K}^*) \neq \emptyset.
\]

We now prove a theorem that dualizes the criteria in Proposition 4. But first, we need a lemma. Note that choice of \( e \) and \( e^* \) is different from the one suggested in Lemma 2.

**Lemma 7.** Let \( \mathcal{K} = \mathcal{K}^1 \times \mathcal{K}^2 \) be a closed convex cone, such that \( \mathcal{K}^2 \) is polyhedral. Consider the pair \( (P_{\mathcal{K}}) \) and \( (D_{\mathcal{K}}) \) with \( e \in (\text{ri } \mathcal{K}^1) \times \{0\} \) and \( e^* \in \text{ri } \mathcal{K}^* \). The following properties hold.

i. \( (P_{\mathcal{K}}) \) has a relative interior feasible solution and \( (D_{\mathcal{K}}) \) satisfies the PPS condition. In particular, the duality gap is zero and both \( (D_{\mathcal{K}}) \) and \( (P_{\mathcal{K}}) \) attain the optimal value.

ii. Let \( (y_1, y_2, y_3) \) be a dual feasible solution satisfying \( y_2 > 0 \). Then, \( s = c - A^T \frac{y_2}{y_1} \) satisfies \( s_1 \in \text{ri } \mathcal{K}^1 \).

iii. If the optimal value is zero, we have either: a) \( \langle c, x^* \rangle < 0 \) and \( F_{D_{\mathcal{K}}}^\text{D} = \emptyset \), or b) \( \langle c, x^* \rangle = 0 \) and \( F_{D_{\mathcal{K}}}^\text{D} \subseteq \mathcal{K} \cap \{x^*\}^\perp \subseteq \mathcal{K} \). In the latter case, we have \( x_1^* \notin \mathcal{K}^1 \) and \( (\mathcal{K}^1)^\perp \).

iv. Let \( (x^*, t^*, w^*) \) be a primal optimal solution. The optimal value is zero if and only if the PPS condition is not satisfied for \( (D_{\mathcal{K}}) \).

**Proof.** i. Let \( t = \frac{1}{(c,e)+1}, \quad w = \frac{1}{(c,e)+1}, \quad \text{and} \quad x = \frac{e^*}{(c,e)+1} \). Then \( (x,t,w) \) is a relative interior solution to \( (P_{\mathcal{K}}) \). Due to the choice of \( e \), \( (0,-1,0) \) is a dual feasible solution such that the associated slack \( ((e,0),1,1) \) belongs to \((\text{ri } \mathcal{K}^1) \times \{0\} \times \text{ri } \mathcal{K}^+ \times \text{ri } \mathcal{K}^+ \). We can then invoke Proposition 4 which ensures that the duality gap is zero and both problems are attained.

ii. Recall that for any closed convex cone \( \mathcal{K} \) we have \( \mathcal{K} + \text{ri } \mathcal{K} = \text{ri } \mathcal{K} \). Hence, \( (c - y_1 e + A^T \frac{y_2}{y_1}) + y_1 e = c - A^T \frac{w}{y_1} \in (\text{ri } \mathcal{K}^1) \times \mathcal{K}^2 \) due to the choice of \( e \) and the fact that \( y_1 > 0 \).

iii. Suppose that the optimal value is zero and let \( (x^*,0,w^*) \) be a primal optimal solution. We must have \( A x^* = 0 \) and \( (c,x^*) = w^* \geq 0 \) and \( x^* \in \mathcal{K}^* \). If \( (c,x^*) < 0 \), then we are done since this is alternative a). Note that this implies the infeasibility of \( (D_{\mathcal{K}}) \), hence \( F_{D_{\mathcal{K}}}^\text{D} = \emptyset \).

On the other hand, if \( (c,x^*) = 0 \), then equation (1) implies \( w^* = 0 \). By equation (2), we have \( (c,x^*) = 1 \), which implies that \( x^* = (x_1^*,x_2^*) \) is such that \( x_1^* \notin \mathcal{K}^1 \). Therefore, there is at least one element \( v \) in \( \mathcal{K} \) for which \( (v,x^*) > 0 \). Hence, \( \mathcal{K} \cap \{x^*\}^\perp \subseteq \mathcal{K} \) and the inclusion is indeed strict. Furthermore, since \( A x^* = 0 \) and \( (c,x^*) = 0 \), we also have \( F_{D_{\mathcal{K}}}^\text{D} \subseteq \mathcal{K} \cap \{x^*\}^\perp \). This is alternative b).

iv. Suppose that the PPS condition is not satisfied. If \( t^* > 0 \), then \( y_2^* = t^* > 0 \) for some dual optimal solution \( (y_1^*,y_2^*,y_3^*) \), because the duality gap is zero. It follows from item ii. that the PPS condition is satisfied, which is impossible.

Conversely, if \( t^* = 0 \) and \( (x^*,0,w^*) \) is an optimal solution for \( (P_{\mathcal{K}}) \), then either a) or b) of item iii. is satisfied. If a) is satisfied, then \( (D_{\mathcal{K}}) \) is infeasible and we are done. If b) is satisfied, then \( (c,x^*) = 0 \), \( A x^* = 0 \) and \( x_1^* \notin \mathcal{K}^1 \). If \( (s_1,s_2) \) is a feasible slack for \( (D_{\mathcal{K}}) \), we have \( \langle s_1,x_1 \rangle + \langle s_2,x_2 \rangle = 0 \). As \( x_1^* \notin \mathcal{K}^1 \), we have that \( s_1 \notin \text{ri } \mathcal{K}^1 \), so \( (D_{\mathcal{K}}) \) cannot possibly satisfy the PPS condition.

**Theorem 8.** Let \( c \in \mathbb{R}^n, \mathcal{L} \subseteq \mathbb{R}^n \) be a subspace and \( \mathcal{K} = \mathcal{K}^1 \times \mathcal{K}^2 \) be a closed convex cone, such that \( \mathcal{K}^2 \) is polyhedral. Then \( (\mathcal{L} + c) \cap (\text{ri } \mathcal{K}^1) \times \mathcal{K}^2 \neq \emptyset \) if and only if one of the conditions below holds:

a) there exists \( x \in \mathcal{K}^* \cap \mathcal{L}^\perp \) such that \( \langle c,x \rangle < 0 \);

b) there exists \( x = (x_1,x_2) \in \mathcal{K}^* \cap \mathcal{L}^\perp \cap \{e\}^\perp \) such that \( x_1 \notin \mathcal{K}^1 \).
Let \( x \in \mathbb{R}^n \times \mathbb{L}_2 \) and \( c = (0, c_2) \) is such that \( (\mathbb{L}_2 + c_2) \cap \mathbb{K}^2 = \emptyset \), then any \( x \in \mathbb{L}^\perp \) must have \( x_1 = 0 \), which implies \( x_1 \in (\mathbb{K}^1)^\perp \). For comparison, we also state the alternative theorem for the condition \( "(\mathbb{L} + c) \cap \ri \mathbb{K} \neq \emptyset" \), which is a consequence of Theorem 8. See also Lemma 3.2 in [25].

Theorem 9. Let \( c \in \mathbb{R}^n \), \( \mathbb{L} \subseteq \mathbb{R}^n \) be a subspace and \( \mathbb{K} \) be a closed convex cone. Then \( (\mathbb{L} + c) \cap \ri \mathbb{K} = \emptyset \) if and only if one of the conditions below holds:

a) there exists \( x \in \mathbb{K}^* \cap \mathbb{L}^\perp \) such that \( (c, x) < 0 \);

b) there exists \( x \in \mathbb{K}^* \cap \mathbb{L}^\perp \cap \{c\}^\perp \) such that \( x \notin \text{lin} \{\mathbb{K}^*\} = \mathbb{K}^\perp \).

5 Distance to polyhedrality and FRA-Poly

One of the goals of facial reduction is to restore strong duality. Here we will discuss FRA-Poly, which is a modification of the facial reduction algorithm that in many cases requires fewer reduction steps than the usual FRA approach. The procedure will be divided in two phases. The first detects infeasibility and restores strong duality, while the second finds the minimal face.

The idea behind the classical FRA, is that whenever strong feasibility fails, we can obtain reducing directions until strong feasibility is satisfied again. Similarly, Phase 1 of FRA-Poly is based on the fact that whenever the generalized condition in Proposition 4 (PPS) fails, we may also obtain reducing directions until the PPS is satisfied, thanks to Theorem 8. In addition, those directions can be found by using \( \{P_\mathbb{K}\} \) and \( \{D_\mathbb{K}\} \) in an appropriate manner, as indicated in Lemma 7. After that, a single extra facial reduction step is enough to go to the minimal face. As the PPS condition is weaker than full-on strong feasibility, FRA-poly has better worst case bounds in many cases.

We now present a disclaimer of sorts. The theoretical results presented in this section and the next stand whether FRA-poly is doable or not for a given \( \mathbb{K} \). If we wish to do facial reduction concretely (even if it is by hand!), we need to make a few assumptions on our computational capabilities and on our knowledge on the lattice of faces of \( \mathbb{K} \). First of all, we must be able to solve problems over faces of \( \mathbb{K} \) such that both the primal and the dual satisfy the PPS condition and we must also be able to do basic linear algebraic operations. Also, for each face \( \mathcal{F} \) of \( \mathbb{K} \) we must know:
1. span $\mathcal{F}$,

2. at least one point $e \in \text{ri} \mathcal{F}$,

3. at least one point $e^* \in \text{ri} \mathcal{F}^*$,

4. whether $\mathcal{F}$ is polyhedral or not.

We remark that apart from knowledge about the polyhedral faces, our assumptions are not very different from what it is usually assumed implicitly in the FRA literature. For symmetric cones, which include direct products of $S^n$, $Q^n$ and $\mathbb{R}^n$, they are reasonable since their lattice of faces is well-understood and every face is again a symmetric cone. So, for instance, $e$ can be taken as the identity element for the corresponding Jordan algebra. On the other hand, if $\mathcal{K}$ is, say, the copositive cone $C^n$, we might have some trouble fulfilling the requirements, inasmuch as our knowledge of the faces of $C^n$ is still lacking.

5.1 FRA-Poly

Henceforth, we will assume that $\mathcal{K}$ is the product of $r$ cones and we will write $\mathcal{K} = \mathcal{K}_1 \times \ldots \times \mathcal{K}_r$. Consider the following FRA variant, which we call FRA-poly.

**Facial Reduction Poly - Phase 1**

**Input:** $[D]$

**Output:** A set of reducing directions $\{d_1, \ldots, d_t\}$. If $[D]$ is feasible, it outputs some face $\mathcal{F} \subseteq \mathcal{K}$ for which the PPS condition holds, together with a dual slack $s'$ for which $s'_j \in \text{ri} \mathcal{F}^j$ for every $j$ such that $\mathcal{F}^j$ is nonpolyhedral. If $[D]$ is infeasible, the directions form a certificate of infeasibility.

1. $\mathcal{F}_1 \leftarrow \mathcal{K}, i \leftarrow 1$

2. Let $e$ be such that $e_j = 0$ if $\mathcal{F}^j_i$ is polyhedral and $e_j \in \text{ri} \mathcal{F}^j_i$ otherwise. Let $e^* \in \text{ri} \mathcal{F}^*_i$. Solve $[P_K]$ and $[D_K]$ with this choice of $e, e^*$ and with $\mathcal{F}_i$ in place of $\mathcal{K}$ to obtain primal dual pairs of optimal solutions $(x^*, t^*, w^*)$ and $(y^*_1, y^*_2, y^*_3)$.

3. If $t^* = 0$ and $\langle e, x^* \rangle < 0$, let $\mathcal{F}_{\min}^D \leftarrow \emptyset$ and stop. $[D]$ is infeasible.

4. If $t^* = 0$ and $\langle e, x^* \rangle = 0$, let $d_i \leftarrow x^*, \mathcal{F}_{i+1} \leftarrow \mathcal{F}_{i} \cap \{d_i\}^\perp, i \leftarrow i + 1$ and return to 2).

5. If $t^* > 0$, $s' \leftarrow c - A^T w^*_3, \mathcal{F} \leftarrow \mathcal{F}_i$ and stop.

Note that Phase 1 of FRA-poly might not end at the minimal face. Nevertheless, Proposition 10 states that, in order to have strong duality, it is enough to have a feasible point $s$ such that $s_i \in \text{ri} \mathcal{K}^i$ for every $i$ that correspond to a nonpolyhedral cone. First, we will show that the output of FRA-Poly is correct.

**Proposition 10.** Phase 1 of FRA-Poly finishes after finding a finite number of directions and its output is correct.

i. if $[D]$ is feasible, then the output face $\mathcal{F}$ contains the minimal face $\mathcal{F}_{\min}^D$ and is such that the PPS condition is satisfied if $\mathcal{K}$ is substituted by $\mathcal{F}$. Moreover, $s'$ is a dual feasible slack such that $s'_j \in \text{ri} \mathcal{F}^j$ for every $j$ such that $\mathcal{F}^j$ is nonpolyhedral.

ii. $[D]$ is infeasible if and only if Step 3. is reached.

**Proof.** The correctness of FRA-poly is guaranteed by Theorem 8 and Lemma 7. In particular, due to Lemma 7, we will be able to find reducing directions by invoking $[P_K]$ and $[D_K]$ with the choice of $e$ and $e^*$ described in Step 2. And if the PPS condition is indeed satisfied, we will know because the optimal value will be positive. In that case, due to item ii. of Lemma 7, we have $s'_j \in \text{ri} \mathcal{F}^j$ for all $j$ such that $\mathcal{F}^j$ is nonpolyhedral.

If the optimal value is zero, we can extract a reducing direction from $[P_K]$ that satisfy alternatives a) or b) of Theorem 8. If it is the former, we will hit Step 3, which implies infeasibility. If it is the latter, we will
have $F_{i+1} \subseteq F_i$, since at least one component of $x^*$ does not belong to the orthogonal complement of the nonpolyhedral part of $F_i$. As $x^*$ is a bona fide reducing direction, we have $F^D_{\text{min}} \subseteq F_{i+1}$. Since $K$ has no infinite descending chain of faces, eventually either Step 5. will be reached or alternative $a)$ of Theorem 8 will hold, which implies that Step 3. will be reached.

As remarked before, Phase 1 of FRA-poly correctly detects infeasibility and if the problem is feasible, we will end up with a face $F$ such that if we reformulate $[\mathcal{D}]$ as a problem over $F$, the duality gap will be zero and the primal will be attained if the common optimal is finite. We will also obtain a solution $s'$ such that it is almost a relative interior point, except for the polyhedral blocks. This means the output face $F$ is such that $F^j = (F^D_{\text{min}})^j$ for every $j$ such that $F^j$ is nonpolyhedral. The next step is showing that we can jump directly to the minimal face in a single facial reduction step.

In Phase 2, we also perform a facial reduction step, but with an important difference. This time, when considering the problems $[P_K]$ and $[D_K]$, instead of using $F$, we relax the nonpolyhedral blocks of $F$ to their span. By doing so, $P_K$ and $D_K$ become polyhedral problems thus ensuring the existence of strict complementary optimal solutions which can easily be found by solving a linear program. From $[P_K]$ we will extract a reducing direction that will allow us to find $F^D_{\text{min}}$ at once. And from $[D_K]$ we will extract a slack $\hat{s}$ for $[\mathcal{D}]$ such that it is a relative interior point of the polyhedral part, but may violate other nonlinear constraints. Recall that $s'$ is a feasible slack for $[\mathcal{D}]$ that is also a relative interior point for the nonpolyhedral part. We then take a strict convex combination of $\hat{s}$ and $s'$ putting a larger weight to $s'$ to tilt $\hat{s}$ towards the relative interior of the nonpolyhedral part and, at the same time, restoring feasibility. As the convex combination will be strict, this will also shift the point towards the relative interior of the polyhedral part as well. Therefore, the resulting point will indeed be a point in $\text{ri } F^D_{\text{min}}$.

**[Facial Reduction Poly - Phase 2]**

**Input:** $\mathcal{D}$, the output of Phase 1: $F$ and $s'$, with $F \neq \emptyset$.

**Output:** $F^D_{\text{min}}$ and dual feasible slack $\hat{s} \in \text{ri } F^D_{\text{min}}$. If $F \neq F^D_{\text{min}}$ then the procedure outputs an extra reducing direction $d$.

1. Let $\hat{K} = \hat{K}^1 \times \ldots \times \hat{K}^r$ such that $\hat{K}^j = F^j$ if $F^j$ is polyhedral and $\hat{K}^j = \text{span } F^j$ otherwise. Let $e \in \text{ri } \hat{K}$ and $e^* \in \text{ri } \hat{K}^*$. Build the systems $[P_K]$ and $[D_K]$.

2. Solve the linear programs $[P_K]$ and $[D_K]$ and obtain a primal and dual pair of strictly complementary optimal solutions $(x^*, t^*, w^*)$ and $(y^*_1, y^*_2, y^*_3)$.

3. If $t^* = 0$, let $d \leftarrow x^*$, $F^D_{\text{min}} \leftarrow F \cap \{x^*\}^\perp$. Let $\hat{s} = c - A^T y^*_1/y^*_1$. Then, we let $\hat{s}$ be a convex combination of $\hat{s}$ and $s'$ such that $\hat{s} \in \text{ri } F^D_{\text{min}}$ and stop.

4. If $t^* > 0$, $F^D_{\text{min}} \leftarrow F$. Let $\hat{s} = c - A^T y^*_1/y^*_1$. Then, we let $\hat{s}$ be a convex combination of $\hat{s}$ and $s'$ such that $\hat{s} \in \text{ri } F^D_{\text{min}}$ and stop.

We now prove that Phase 2 of FRA-Poly is correct. This is a consequence of the following two results. First, recall that $[\mathcal{D}]$ is strongly infeasible if and only if there is $x$ such that $x \in K^* \cap \ker A$ and $\langle c, x \rangle < 0$, see, for instance, Lemma 5 in [14].

**Theorem 11.** Consider the pair of problems $[P_K]$ and $[D_K]$, under the setting of Lemma 3. If there is a pair of dual strict complementary solutions $(x^*, t^*, w^*)$, $(y^*_1, y^*_2, y^*_3)$ (see comments after Proposition 1) and the optimal value is zero, then $[\mathcal{D}]$ is either strongly infeasible or $F^D_{\text{min}} = K \cap \{x^*\}^\perp$. In the latter case, it also holds that $c - A^T y^*_1/y^*_1 \in \text{ri } F^D_{\text{min}}$.

Proof. If $[\mathcal{D}]$ is not strongly infeasible, then we must have $t^* = w^* = 0$. Then, dual strict complementarity implies that the third inequality of $[D_K]$ must be strict, that is, $y^*_1 > 0$. Also due to dual strict complementarity we have that $cy^*_1 - A^T y^*_2 \in \text{ri } (K \cap \{x^*\}^\perp)$. Since $y^*_1 > 0$, we have $c - A^T y^*_1/y^*_1 \in \text{ri } (K \cap \{x^*\}^\perp)$, which shows that $F^D_{\text{min}} = K \cap \{x^*\}^\perp$. □
In particular, when \( K \) is polyhedral, both \( [D_K] \) and \( [P_K] \) are polyhedral problems. In this case, strict complementary solutions are ensured to exist, which is a consequence of Goldman-Tucker Theorem. This can also be inferred by the the results of McLinden \[15\] and Akgül \[1\]. We also remark that a strict complementary solution of a polyhedral problem can be found by solving a single linear program, see, for instance, the article by Freund, Roundy and Todd \[7\] and the related work by Mehrotra and Ye \[16\].

**Theorem 12.** The output of Phase 2 of FRA-poly is correct.

**Proof.** Our assumption here is that the outputs of Phase 1 are such that \( F \neq \emptyset \) and \( s' \) satisfies \( s'_j \in \text{ri} F^j \) for every \( j \) such that \( F^j \) is nonpolyhedral.

First, consider the case where \( F \) is not the minimal face, i.e., \( (\text{ri} F) \cap (c + \text{range} A^T) = \emptyset \). By Theorem 8, this happens if and only if \( F \) and \( c + \text{range} A^T \) can be properly separated. Therefore, there exists \( x \in F^* \) such that \( Ax = 0 \) and either: i) \( \langle c, x \rangle < 0 \), or ii) \( \langle c, x \rangle = 0 \) and \( x \notin F^\bot \). As infeasibility is detected at Phase 1, alternative i) cannot occur and any \( x \) inducing proper separation between \( F \) and \( c + \text{range} A^T \) must satisfy alternative ii). Note that \( x \) is a reducing direction as well.

However, because \( F \) is the output of Phase 1 of FRA-Poly, we have that neither alternative a) nor alternative b) of Theorem 8 can hold. So the possible reducing directions \( x \) must be such that \( x_j \in \text{lin} ((F^j)^*) = (F^j)^\bot = (\text{span} F^j)^\bot \) for every \( j \) such that \( F^j \) is not polyhedral, lest we run afoul of Theorem 8. We can then conclude that the possible reducing directions are confined to the polyhedral cone \( \hat{K}^j \), where \( \hat{K} = \hat{K}^1 \times \ldots \times \hat{K}^r \) is such that \( K^j = F^j \) if \( F^j \) is polyhedral and \( \hat{K}^j = \text{span} F^j \) otherwise. This is precisely the cone appearing in Phase 2 of FRA-poly.

If we build the systems \( [F_K] \) and \( [D_K] \) precisely as in Lemma 2 using \( \hat{K} \), we will obtain a pair of linear programs. Therefore they have a pair of strictly complementary optimal solutions \( (x^*, t^*, w^*) \), \( (y_1^*, y_2^*, y_3^*) \). Because \( F \) is not the minimal face, we have \( t^* = 0 \), so we are under the hypothesis of Theorem 11. Let \( \tilde{s} = c - A^T y_3^*/y_1^* \).

We will prove that \( F^D_{\text{min}} = F \cap \{x^*\}^\bot \) and that some convex combination of \( s' \) and \( \tilde{s} \) is a relative interior feasible solution of \( [D] \). Let \( z_\beta = \beta s' + (1 - \beta) \tilde{s} \). For all \( \beta \in (0, 1) \) and all \( j \) such that \( F^j \) is polyhedral, we have \( z_\beta j \in \text{ri} (F^j \cap \{x_j^*\})^\bot \), because \( \tilde{s} j \in \text{ri} (F^j \cap \{x_j^*\})^\bot \) and \( s' \) is feasible. If \( F^j \) is not polyhedral, then \( F^j \cap \{x_j^*\}^\bot = \emptyset \), since \( x_j \in (F^j)^\bot \). Because \( \tilde{s} j \in \text{span} F^j \) and \( s'_j \in \text{ri} F^j \), for \( \beta \) sufficiently close to 1 we have \( z_\beta j \in \text{ri} F^j \). Therefore, it is possible to select \( \beta \in (0, 1) \) such that \( z_\beta j \in \text{ri} (F^j \cap \{x_j^*\})^\bot \) for all \( j \).

This also shows that \( F^D_{\text{min}} = F \cap \{x^*\}^\bot \).

If \( F \) was already the minimal face to begin with, then \( t^* > 0 \). We can then proceed in a similar fashion. The only difference is that due to \( [3] \), we will have that \( \tilde{s} = c - A^T y_2^*/y_1^* \) satisfies \( \tilde{s} j \in \text{ri} (F^j) \) for every \( j \) such that \( F^j \) is polyhedral. And as before, we can select a convex combination of \( s' \) and \( \tilde{s} \) belonging to the relative interior of \( F^D_{\text{min}} \).

\[ \square \]

### 5.2 Distance to Polyhedrality

In order to bound the number of directions obtained through FRA-poly, we introduce the notion of distance to polyhedrality. In what follows, if we have a chain of faces \( F_1 \sqsubseteq \ldots \sqsubseteq F_\ell \), the length of the chain is defined to be \( \ell \).

**Definition 13.** Let \( K \) be a nonempty closed convex cone. The distance to polyhedrality is the length minus one of the longest strictly ascending chain of nonempty faces \( F_1 \sqsubseteq \ldots \sqsubseteq F_\ell \) which satisfies:

1. \( F_1 \) is polyhedral;
2. \( F_j \) is not polyhedral for \( j > 1 \).

We will denote the distance to polyhedrality by \( \ell_{\text{poly}}(K) \).

The distance to polyhedrality is a well-defined concept, because the lineality of \( K \) is always an exposed polyhedral face of \( K \). Therefore it is always possible to consider the chain \( \text{lin}(K) \). Moreover, \( \ell_{\text{poly}}(K) \) counts...
the maximum number of facial reduction steps that can be taken before we reach a polyhedral face. Therefore a necessary and sufficient condition for a cone to be polyhedral is that $\ell_{\text{poly}}(K) = 0$.

**Example 14.** See section 2 and examples 2.5 and 2.6 in [18] for more details on the facial structure of $S^n_+$ and $Q^n$. For the positive semidefinite cone $S^n_+$, we have $\ell_{\text{poly}}(S^n_+) = n - 1$. For a single Lorentz cone $Q^n = \{(x, \lambda) \in \mathbb{R} \times \mathbb{R}^{n-1} \mid x_0 \geq |\lambda|_2 \}$, we have $\ell_{\text{poly}}(Q^n) = 1$ if $n > 2$. This is because the Lorentz cone only has three types of faces: $(0)$, $Q^n$ or the half-lines running along the boundary. For comparison, the longest chain of nonempty faces of $S^n_+$ has length $n + 1$ and the one for $Q^n$ has length $3$.

**Proposition 15.** Let $K = K^1 \times \ldots \times K^r$. If $[D]$ is feasible, Phase 1 of FRA-Poly stops after finding at most $\sum_{i=1}^r \ell_{\text{poly}}(K^i)$ directions.

If $[D]$ is infeasible, Phase 1 stops after finding at most $1 + \sum_{i=1}^r \ell_{\text{poly}}(K^i)$ directions.

**Proof.** Due to the choice made at Step 2 of Phase 1 of FRA-Poly and the analysis done in Theorem 8 whenever $(c, x^*) = 0$, we have that $x_j^* \notin (F_I^j)^\perp$ for at least one nonpolyhedral cone $F_I^j$. This means that $F_{I+1}^j$ is a proper face of $F_I^j$. In other words, whenever a new proper face is found, it is because we are making progress towards a polyhedral face for at least one nonpolyhedral cone. Therefore, after finding $\ell = \sum_{i=1}^r \ell_{\text{poly}}(K^i)$ directions, $F_{I+1}^j$ is polyhedral.

We can now consider the immediate corollary.

**Corollary 16.** Let $K = K^1 \times \ldots \times K^r$. The minimum face $F_{\text{min}}^D$ that contains the feasible region of $[D]$ can be found in no more than $1 + \sum_{i=1}^r \ell_{\text{poly}}(K^i)$ facial reduction steps.

**Proof.** If $[D]$ is infeasible, then $F_{\text{min}}^D = \emptyset$ and the result follows from Proposition 15. So suppose now that $[D]$ is feasible. Then FRA-Poly ends after finding at most $\sum_{i=1}^r \ell_{\text{poly}}(K^i)$. Due to Theorem 12 at most one extra direction is needed to jump to the minimal face.

The number of directions found in FRA-poly can also be bounded by a quantity that depends on $L = \text{range} \mathcal{A}^T$ and $c$. First, note that whenever FRA-Poly hits Step 4, we have that $d_i$ does not belong to the span of the previous directions $\{d_1, \ldots, d_{i-1}\}$. Otherwise, we would have $K \cap \{d_1\}^\perp \cap \ldots \cap \{d_{i-1}\}^\perp = K \cap \{d_1\}^\perp \cap \ldots \cap \{d_{i-1}\}^\perp \cap \{d_i\}^\perp$, which would contradict the fact that $F_i \subseteq F_{i+1}$. If $[D]$ is feasible then all the directions are contained $L^\perp \cap \{c\}^\perp$, so it follows that $\ell \leq \text{dim } L + \text{dim } \{c\}^\perp$. If $[D]$ is infeasible, all the directions except $d_i$ are in $L^\perp \cap \{c\}^\perp$, so we have $\ell \leq \text{dim } L \cap \{c\} + 1$.

Recall that $\ell_K$ is the length of the longest chain of strictly ascending nonempty faces of $K$. If one uses the “classical” facial reduction approach, it takes no more than $\ell_K - 1$ facial reduction steps to find the minimal face, when $[D]$ is feasible. See, for instance, Theorem 1 in [19] or Corollary 3.1 in [25]. If $[D]$ is infeasible, an extra direction might be needed, which is the one that will hit Step 3 in the Generic Facial Reduction of Section 3. When $K$ is a direct product of several cones, we have $\ell_K = 1 + \sum_{i=1}^r (\ell_{K^i} - 1)$. We will end this subsection by showing that, under the relatively weak hypothesis that $K^i$ is not a subspace, we have $\ell_{\text{poly}}(K^i) < \ell_{K^i} - 1$. This means that the number of steps needed in FRA-Poly is no worse than the classical FRA and if we have the direct product of at least two cones that are not subspaces, FRA-Poly compares favorably to the classical approach. First, we need the following technical lemma.

**Lemma 17.** Suppose that $K$ is pointed, that is, $\text{lin}(K) = \{0\}$ and that its dimension is greater than zero, then $\ell_{\text{poly}}(K) < \ell_K - 1$. 

12
We now substitute the hypothesis of pointedness by the weaker assumption that $\mathcal{K}$ is not a subspace.

**Theorem 18.** If $\mathcal{K}$ is not a subspace then $\ell_{\text{poly}}(\mathcal{K}) < \ell_{\mathcal{K}} - 1$. In particular, if $\mathcal{K}$ is the direct product of $r$ closed convex cones that are not subspaces we have:

$$r + 1 + \sum_{i=1}^{r} \ell_{\text{poly}}(\mathcal{K}^i) \leq 1 + \sum_{i=1}^{r} (\ell_{\mathcal{K}^i} - 1).$$

**Proof.** Let $U = \text{lin} \mathcal{K}$. Then we have $\mathcal{K} = (\mathcal{K} \cap U^\perp) + U$. If we let $\hat{K} = \mathcal{K} \cap (U^\perp)$, we have that $\text{lin} (\hat{K}) = \{0\}$. It can be shown that there is a bijection between the faces of $\mathcal{K}$ and the set $\{F + U \mid F$ is a face of $\hat{K}\}$. There is also a correspondence between the polyhedral faces of $\mathcal{K}$ and the set $\{F + U \mid F$ is a polyhedral face of $\hat{K}\}$. The assumption that $\mathcal{K}$ is not a subspace implies that the dimension of both $\mathcal{K}$ and $\hat{K}$ is greater than zero. As $\ell_{\mathcal{K}} = \ell_{\hat{K}}$, the result follows from applying Lemma 17 to $\mathcal{K}$. \qed

of Lemma 17. Let $e^* \in \text{ri} \mathcal{K}^*$, then the set $C = \{x \in \mathcal{K} \mid \langle x, e^* \rangle = 1\}$ is compact. This is because the recession cone of $C$ consists of the elements in $\mathcal{K}$ which are orthogonal to $e^*$, but pointedness imply that $0$ is the only element meeting these criteria. Now, the Krein-Milman Theorem implies that a nonempty compact convex set has at least one extreme point $z$. Then, one can verify that the half-line $h_z = \{\alpha z \mid \alpha \geq 0\}$ is an one-dimensional face of $\mathcal{K}$. Let $F$ be a face of $\mathcal{K}$ with dimension greater than zero. We can apply the same argument to conclude that $F$ has at least one extreme ray, i.e., a face of dimension one.

So $\ell_{\mathcal{K}} \geq 2$, since we have the chain $\{0\} \subsetneq h_z$. If $\mathcal{K}$ is polyhedral, we are done, since $\ell_{\text{poly}}(\mathcal{K}) = 0$. We now move on to the nonpolyhedral case. Let $F_1 \subsetneq \ldots \subsetneq F_\ell$ be a strictly ascending chain of faces such that $F_1$ is polyhedral, $F_\ell$ is not polyhedral for $j > 1$ and $\ell - 1 = \ell_{\text{poly}}(\mathcal{K})$. Due to non-polyhedrality, $\ell \geq 2$. We now consider two cases. If the dimension of $F_1$ is greater or equal than one, we can augment the chain by adding the face $\{0\}$ at the beginning. In this case, we have $\ell_{\mathcal{K}} - 1 \geq \ell > \ell_{\text{poly}}(\mathcal{K})$.

On the other hand, if $F_1 = \{0\}$, we have for sure that $F_2$ has dimension greater or equal than two. This is because $F_2$ is not polyhedral, so it cannot be an extreme ray. However, due to the previous argument, $F_2$ has at least one extreme ray $h_z$, so we can augment the chain by inserting $h_z$ between $F_1$ and $F_2$. This also shows that $\ell_{\mathcal{K}} - 1 \geq \ell > \ell_{\text{poly}}(\mathcal{K})$. \qed

6 Applications of FRA-Poly

In this section, we discuss applications of FRA-poly. In the first one, we sharpen a result proven by Liu and Pataki on the geometry of weakly infeasible problems. In the second, we show that the singularity degree of problems the doubly nonnegative cone is at most $n + 1$. We need two definitions first. We define the singularity degree $d(D)$ of $[D]$ as the minimum number of facial reduction steps needed to find $F_{\text{min}}^D$. No matter the facial reduction strategy used, one must identify at least $d(D)$ directions before $F_{\text{min}}^D$ is reachable.

**Definition 19 (Singularity degree).** Consider the set of possible outputs $\{d_1, \ldots, d_\ell\}$ of the Generic Facial Reduction algorithm in Section 3. The singularity degree of $[D]$ is the minimum $\ell$ among all the possible outputs and is denoted by $d(D)$.

As far as we know, the expression “singularity degree” in this context is due to Sturm in [23], where he showed the connection between the singularity degree of a positive semidefinite program and a Hölderian error bound. In the recent work by Liu and Pataki [10], there is also a definition of singularity degree for general linear conic problems, see Definition 6 therein. One difference, however, is that Liu and Pataki only define the singularity degree for feasible problems and, indeed, when $[D]$ is feasible, their definition matches with Definition 19.\footnote{To see that, first note that every nonzero element in $\mathcal{K}$ can be written as a product $\alpha x$ with $x \in C$ and $\alpha > 0$. It is enough to consider the case where $\beta \alpha_1 x + (1 - \beta)\alpha_2 y = \gamma z$ with $\alpha_1, \alpha_2, \gamma > 0$, $x, y \in C$ and $\beta \in (0, 1)$. Then $\frac{\beta \alpha_1}{\gamma} + \frac{(1 - \beta)\alpha_2}{\gamma} = 1$. Since $z$ is an extreme point, we have $x = y = z$. Therefore, $\alpha_1 x, \alpha_2 y \in h_z$.}
The singularity degree of $[D]$ is a quantity that depends on $c, A$ and $K$. However, it is possible to give uniform bounds for $d(D)$ that do not depend on $c, A$. For example, the classical facial reduction strategy gives the bounds $d(D) \leq \ell_K - 1$ when $[D]$ is feasible and $d(D) \leq \ell_K$ when $[D]$ is infeasible. Corollary 16 readily implies that $d(D) \leq 1 + \sum_{i=1}^r \ell_{poly}(K_i^*)$, no matter whether $[D]$ is feasible or not. Due to Theorem 18 this bound is likely to compare favorably to $\ell_K - 1 = \sum_{i=1}^r (\ell_{K_i^*} - 1)$.

As mentioned before, the singularity degree only depends on $c, A$ and $K$. Finding the minimal face $F_{\text{min}}$ ensures that no matter which $b$ we select, as long as the problem is bounded, there will be zero duality gap and primal attainment. This suggests the following definition that also depends on $b$ and, thus, produce a less conservative quantity.

**Definition 20** (Distance to strong duality). The distance to strong duality $d_{\text{str}}(D)$ is the minimum number of facial reduction steps (at $[D]$) needed to ensure $\theta_P = \theta_D$, where $(\hat{P})$ is the problem $\inf \{ \langle c, x \rangle \mid Ax = b, x \in F_{\ell+1} \}$ and $F_{\ell+1}$ is obtained after a sequence of facial reduction steps. If $-\infty < \theta_D < +\infty$, we also require attainment of $\theta_P$.

Similarly, we define $d_{\text{str}}(P)$ as the minimum number of facial reduction steps needed to ensure that $\theta_P = \theta_D$ and that $\theta_D$ is attained when $-\infty < \theta_P < +\infty$. It is understood that $(\hat{D})$ is the problem in dual standard form arising after some sequence of facial reduction steps is done at $(\hat{P})$.

Clearly, we have $d_{\text{str}}(D) \leq d(D)$. However, since Phase 1 of FRA-Poly restores strong duality in the sense of Definition 20, we obtain the nontrivial bound $d_{\text{str}}(D) \leq \sum_{i=1}^r \ell_{poly}(K_i^*)$.

### 6.1 Weak infeasibility

Let $V$ denote the affine space $c + \text{range } A^T$ and let the tuple $(V, K)$ denote the feasibility problem of seeking an element in the intersection $V \cap K = F^K$. In [H], we showed that if $K = S^+_n$ and $[D]$ is weakly infeasible, then there is a subaffine space $V'$ contained in $V$ of dimension at most $n - 1$ such that $(V', K)$ is also weakly infeasible. This can be interpreted as saying that “we need at most $n - 1$ directions to approach the positive semidefinite cone”. In [L1], Liu and Pataki generalized this result and proved that those affine spaces always exist and $\ell_{K^*} - 1$ is an upper bound for the dimension of $V'$, see Theorem 9 therein. We proved a bound of $r$ for the direct product of $r$ Lorentz cones [H3], which is tighter than the one in [L1]. Here we will refine these results. We first need an auxiliary result concerning the preservation of strong infeasibility.

**Lemma 21.** Let $d \in \text{range } A^T \cap K$ and let by $\hat{K} = (K^* \cap \{d\}^\perp)^*$. Then $[D]$ is strongly infeasible if and only if $(\hat{D})$ is strongly infeasible, where $(\hat{D})$ is the problem with $K$ in place of $K$.

**Proof.** $(\Rightarrow)$ This part is clear, since $K \subseteq \hat{K}$.

$(\Rightarrow)$. Strong infeasibility of $[D]$ is equivalent to the existence of $x$ such that $x \in K^* \cap \ker A$ and $\langle c, x \rangle < 0$. Since $d \in \text{range } A^T$, we have $\langle x, d \rangle = 0$. By the same principle, $x$ induces strong infeasibility for $(\hat{D})$ as well.

**Theorem 22.**  

1. Let $(P')$ be the optimization problem $\inf \{ \langle c, x \rangle \mid Ax = 0, x \in K^* \}$. Then, $[D]$ is not strongly infeasible if and only if there are:

   a) a sequence of reducing directions $\{d_1, \ldots, d_r\}$ for $(P')$ restoring strong duality in the sense of Definition 20 with $\ell = d_{\text{str}}(P')$;

   b) a feasible slack $\hat{s}$ to $(\hat{D})$, where $(\hat{D})$ is the problem $\sup \{ 0 \mid c - A^T y \in (K^* \cap \{d_1\}^\perp \ldots \cap \{d_r\}^\perp)^* \}$.

2. If $[D]$ is not strongly infeasible, there is an affine subspace $V' \subseteq V$ such that $(V', K)$ is not strongly infeasible and the dimension of $V'$ satisfies

   \[ \dim (V') \leq d_{\text{str}}(P') \leq \sum_{i=1}^r \ell_{poly}(K_i^*) \]

3. If $[D]$ is weakly infeasible, then $(V', K)$ is weakly infeasible as well, where $V'$ is the space of item ii.
Proof. i. ($\Rightarrow$) Consider the optimization problem $\theta_{P'} = \sup \{c - A^Ty \in \mathcal{K}\}$. Its corresponding primal is $\theta_{P'} = \inf \{(c,x) | Ax = 0, x \in \mathcal{K}^+\}$. Due to the assumption that $[D]$ is not strongly infeasible, we have $\theta_{P'} = 0$. Now, let $\{d_1, \ldots, d_\ell\}$ be a sequence of reducing directions for $P'$ that restores strong duality in the sense of Definition 20 with $\ell = d_{str}(P')$. This time, the reducing directions define faces $\mathcal{F}_1 \supseteq \cdots \supseteq \mathcal{F}_{\ell+1}$ of $\mathcal{K}^*$ and we have $\mathcal{F}_{\ell+1} = \mathcal{K}^* \cap \{d_1\}^\perp \cap \cdots \cap \{d_\ell\}^\perp$.

Now $(P')$ is equivalent to $\theta_P = \inf \{(c,x) | Ax = 0, x \in \mathcal{F}_{\ell+1}\}$ and the corresponding dual is $\theta_{\tilde{D}} = \sup \{0 | c - A^T y \in \mathcal{F}_{\ell+1}^*\}$. Since facial reduction done at $(P')$ preserves the primal optimal value, we have $\theta_P = \theta_{P'} = 0$. Due to our assumption on the facial reduction sequence, we have $\theta_{\tilde{D}} = 0$ and $\theta_P$ is attained. It follows that there is $\hat{s} = c - A^T \hat{y}$ such that $\hat{s} \in \mathcal{F}_{\ell+1}^*$.

($\Leftarrow$) If $[D]$ were strongly infeasible, then $\theta_{P'} = -\infty$. As $\theta_{P'} = \theta_P$, it would be impossible for $(\tilde{D})$ to admit a feasible solution.

ii. Let $\mathcal{V}'$ be the affine space $\hat{s} + \mathcal{L}'$, where $\mathcal{L}'$ is spanned by the directions $\{d_1, \ldots, d_\ell\}$ of item i. and $\hat{s}$ is a feasible slack for $(\tilde{D})$. Since $\ell = d_{str}(P')$, we have dim $\mathcal{V}' = d_{str}(P')$. Suppose for the sake of contradiction that $(\mathcal{V}', \mathcal{K})$ is strongly infeasible. Then, we can use the same set $\{d_1, \ldots, d_\ell\}$ as reducing directions for $\inf \{(s,x) | x \in \mathcal{L}', x \in \mathcal{K}^*\}$. However, Lemma 21 implies that $\sup \{0 | s \in \mathcal{V}' \cap \mathcal{F}_{\ell+1}^*\}$ is strongly infeasible. But this is impossible, since $\hat{s}$ is a feasible solution.

Since the number steps required for Phase 1 of FRA-Poly gives an upper bound for $d_{str}(P')$, we obtain $d_{str}(P') \leq \sum_{i=1}^{\ell} \ell_{poly}((\mathcal{K}^*)^*)$.

iii. Finally, when $[D]$ is infeasible, since $\mathcal{V}' \subseteq \mathcal{V}$ and $(\mathcal{V}', \mathcal{K})$ is not strongly infeasible, then it must be the case that $(\mathcal{V}', \mathcal{K})$ is weakly infeasible.

Due to Theorem 18 the bound in Theorem 22 will usually compare favorably to $\ell_{\mathcal{K}^*} - 1$. Moreover, it also recovers the bounds described in [11, 13]. Note also that the problem $(P')$ appearing in Theorem 22 is such that $\theta_{P'} < 0$ if and only if $[D]$ is strongly infeasible.

6.2 An application to the intersection of cones

In this subsection, we discuss the case where $\mathcal{K} = \mathcal{K}^1 \cap \mathcal{K}^2$. We can rewrite $[D]$ as a problem over $\mathcal{K}^1 \times \mathcal{K}^2$ by duplicating the entries:

\[
\sup_y (b, y) \quad (D_{\text{dup}})
\]

subject to \( (c - A^Ty, c - A^Ty) \in \mathcal{K}^1 \times \mathcal{K}^2 \)

\[
\inf_x (c, x^1 + x^2) \quad (P_{\text{dup}})
\]

subject to \( A(x^1 + x^2) = b \)

\( (x^1, x^2) \in (\mathcal{K}^1)^* \times (\mathcal{K}^2)^* \).

While $[D_{\text{dup}}]$ is entirely equivalent to $[D]$, the situation for $[P_{\text{dup}}]$ is subtler. It is true that $\theta_{P_{\text{dup}}} = \theta_P$ and that if $[P_{\text{dup}}]$ is attained, then $[P]$ must be attained. However, the converse is not true and $[P_{\text{dup}}]$ might fail to be attained even if $[P]$ is attained. This situation can happen if $((\mathcal{K}^1)^* + (\mathcal{K}^2)^*) \not\subseteq \mathcal{K}^*$.

Still, if we apply FRA-Poly to $[D_{\text{dup}}]$, we will obtain a face $\mathcal{F}^1 \times \mathcal{F}^2$ of $\mathcal{K}^1 \times \mathcal{K}^2$. Doing facial reduction using the formulation $[D_{\text{dup}}]$ might be more convenient, since we need to search for reducing directions in $(\mathcal{K}^1)^* \times (\mathcal{K}^2)^*$ instead of $\text{cl}((\mathcal{K}^1)^* + (\mathcal{K}^2)^*)$ and deciding membership in $(\mathcal{K}^1)^* \times (\mathcal{K}^2)^*$ could be more straightforward than doing the same for $\text{cl}((\mathcal{K}^1)^* + (\mathcal{K}^2)^*)$. Before we proceed we need an auxiliary result.

If $\mathcal{K} = \mathcal{K}^1 \cap \mathcal{K}^2$, it is always true that the intersection of a face of $\mathcal{K}^1$ with a face of $\mathcal{K}^2$ results in a face of $\mathcal{K}$. However, it is not entirely obvious that every face of $\mathcal{K}$ arises as an intersection of faces of $\mathcal{K}^1$ and $\mathcal{K}^2$, so we remark that as a proposition although it is probably a well-known result.
Proposition 23. Let \( \mathcal{F} \) be a nonempty face of \( \mathcal{K} = \mathcal{K}_1 \cap \mathcal{K}_2 \). Let \( \mathcal{F}_1 \) and \( \mathcal{F}_2 \) be the minimal faces of \( \mathcal{K}_1 \) and \( \mathcal{K}_2 \), respectively, containing \( \mathcal{F} \). Then \( \mathcal{F} = \mathcal{F}_1 \cap \mathcal{F}_2 \) and \( \mathcal{F}^* = (\mathcal{F}_1)^* + (\mathcal{F}_2)^* \).

Proof. We have \( \mathcal{F}_1 \cap \mathcal{F}_2 \supseteq \mathcal{F} \) and we will prove that \( \mathcal{F}_1 \cap \mathcal{F}_2 = \mathcal{F} \). A first observation is that by the choice of \( \mathcal{F}_1 \) and \( \mathcal{F}_2 \), we have \( \mathbf{r}_i(\mathcal{F}) \subseteq \mathbf{r}_i(\mathcal{F}_1) \) and \( \mathbf{r}_i(\mathcal{F}) \subseteq \mathbf{r}_i(\mathcal{F}_2) \). In particular, this implies that \( \mathbf{r}_i(\mathcal{F}_1) \cap \mathbf{r}_i(\mathcal{F}_2) \neq \emptyset \). Therefore, \( \mathbf{r}_i(\mathcal{F}_1 \cap \mathcal{F}_2) = \mathbf{r}_i(\mathcal{F}_1) \cap \mathbf{r}_i(\mathcal{F}_2) \), see Theorem 6 in [22]. We conclude that \( \mathbf{r}_i(\mathcal{F}) \cap \mathbf{r}_i(\mathcal{F}_1) \cap \mathbf{r}_i(\mathcal{F}_2) \neq \emptyset \). Since \( \mathcal{F}_1 \cap \mathcal{F}_2 \) is a face of \( \mathcal{K} \), it follows that \( \mathcal{F} = \mathcal{F}_1 \cap \mathcal{F}_2 \).

Because \( \mathbf{r}_i(\mathcal{F}_1) \cap \mathbf{r}_i(\mathcal{F}_2) \neq \emptyset \), a classical closedness criteria implies that \( (\mathcal{F}_1)^* + (\mathcal{F}_2)^* \) is closed (see Corollary 16.4.2 in [22]), so that \( \mathcal{F}^* = \operatorname{cl}( (\mathcal{F}_1)^* + (\mathcal{F}_2)^* ) = (\mathcal{F}_1)^* + (\mathcal{F}_2)^* \).

\( \square \)

Theorem 24. Let \( \mathcal{K} = \mathcal{K}_1 \cap \mathcal{K}_2 \).

i. Let \( \hat{\mathcal{F}} = \mathcal{F}_1 \times \mathcal{F}_2 \) be the minimal face of \( \mathcal{K}_1 \times \mathcal{K}_2 \) containing the feasible slacks of \( D_{\text{dup}} \). Then, \( \mathcal{F}_{\text{min}} = \mathcal{F}_1 \cap \mathcal{F}_2 \).

ii. The singularity degree of \( [D] \) satisfies \( \operatorname{d}(D) \leq \operatorname{d}(D_{\text{dup}}) \leq 1 + \ell_{\text{poly}}(\mathcal{K}_1) + \ell_{\text{poly}}(\mathcal{K}_2) \).

iii. The distance to strong duality satisfies \( \operatorname{d}_{\text{str}}(D) \leq \operatorname{d}_{\text{str}}(D_{\text{dup}}) \leq \ell_{\text{poly}}(\mathcal{K}_1) + \ell_{\text{poly}}(\mathcal{K}_2) \).

Proof. i. Note that \( \mathcal{F}_1 \) must be the minimal face of \( \mathcal{K}_1 \) containing \( \mathcal{F}_1^* = \{ c - A y \in \mathcal{K} \} \). Because if some proper face \( \hat{\mathcal{F}}_\ell \) of \( \mathcal{F}_1 \) is minimal, then \( \hat{\mathcal{F}} \times \mathcal{F}_2 \) contains the feasible slacks of \( D_{\text{dup}} \), which contradicts the minimality of \( \hat{\mathcal{F}} \). The same must hold for \( \mathcal{F}_2 \). Then Proposition 23 implies \( \mathcal{F}_{\text{min}} = \mathcal{F}_1 \cap \mathcal{F}_2 \).

ii. In order to prove that the singularity degree of \( [D] \) is bounded by \( \operatorname{d}(D_{\text{dup}}) \), we need to check whether a sequence of reducing directions for \( D_{\text{dup}} \) translate into reducing directions for \( [D] \). A sequence of reducing directions \( \{ d_1, \ldots, d_\ell \} \) and corresponding faces for \( [D_{\text{dup}}] \) are such that:

(a) \( \hat{\mathcal{F}}_i = \mathcal{F}_1^* \times \mathcal{F}_2^* \).

(b) \( d_i = (d_i^1, d_i^2) \in \hat{\mathcal{F}}_i^* \), \( \hat{\mathcal{F}}_{i+1}^* = (\hat{\mathcal{F}}_i^*) \cap \mathcal{F}_1 \cap \mathcal{F}_2 \), and \( A(d_i^1 + d_i^2) = 0 \), \( \langle c, d_i^1 + d_i^2 \rangle \leq 0 \) for all \( i \). We will prove that \( \{ d_1^1 + d_2^1, \ldots, d_\ell^1 + d_\ell^2 \} \) form a valid sequence of reducing directions in \( \mathcal{F}_1 \), except that some of the directions might fail to induce a proper face. The only thing missing is to prove that \( d_i^1 + d_i^2 \in \mathcal{F}_i^* \) for every \( i \), with \( \mathcal{F}_{i+1} = \mathcal{F}_i \cap \mathcal{F}_1 \cap \mathcal{F}_2 \). And then, if \( d_i^1 + d_i^2 \in \mathcal{F}_i^* \cap \{ c \} \), we simply discard \( d_i^1 + d_i^2 \).

We will prove by induction that \( \mathcal{F}_i = \hat{\mathcal{F}}_i^* \cap \mathcal{F}_i^* \) for every \( i \) and that \( d_i^1 + d_i^2 \in \mathcal{F}_i^* \) for every \( i \). First, note that \( \mathcal{F}_1 = \hat{\mathcal{F}}_1^* \cap \mathcal{F}_1^* = \mathcal{K} \). Because \( \mathcal{K}^* = \operatorname{cl}( (\mathcal{K}_1)^* + (\mathcal{K}_2)^* ) \), we also have \( d_1^1 + d_1^2 \in \mathcal{K}^* \). This takes care of the basis of induction.

Now, suppose that the statement holds true for some \( i \) and let us show that it holds for \( i + 1 \). We have \( \mathcal{F}_{i+1}^* = \mathcal{F}_i^* \cap \mathcal{F}_1 \cap \mathcal{F}_2 \), and \( A(d_i^1 + d_i^2) = 0 \), \( \langle c, d_i^1 + d_i^2 \rangle \leq 0 \) for all \( i \). We will prove by induction that \( d_i^1 + d_i^2 \in \mathcal{F}_i^* \) for every \( i \), with \( \mathcal{F}_{i+1} = \mathcal{F}_i \cap \mathcal{F}_1 \cap \mathcal{F}_2 \). And then, if \( d_i^1 + d_i^2 \in \mathcal{F}_i^* \cap \{ c \} \), we simply discard \( d_i^1 + d_i^2 \).

We have proved that the chain of faces \( \mathcal{F}_1 \supseteq \ldots \supseteq \mathcal{F}_\ell \) is such that \( \mathcal{F}_1 = \mathcal{K} \) and \( \mathcal{F}_\ell = \mathcal{F}_{\text{min}} \). However, some of the containments might fail to be strict. This poses no problem, since it is enough to remove the reducing directions that provide no decrease. This shows that \( \operatorname{d}(D) \leq \operatorname{d}(D_{\text{dup}}) \). If we apply both Phases of FRA-poly to \( D_{\text{dup}} \), we obtain the bound \( \operatorname{d}(D_{\text{dup}}) \leq 1 + \ell_{\text{poly}}(\mathcal{K}_1) + \ell_{\text{poly}}(\mathcal{K}_2) \).

iii. In the proof of item ii., it was shown that we can obtain reducing directions from \( D \) from reducing directions from \( D_{\text{dup}} \). Suppose that \( \{ d_1, \ldots, d_\ell \} \) restores strong duality for \( D_{\text{dup}} \), in the sense of Definition 20. Then, it is straightforward to check that \( \{ d_1^1 + d_1^2, \ldots, d_\ell^1 + d_\ell^2 \} \) restores strong duality for \( [D] \). This shows that \( \operatorname{d}_{\text{str}}(D) \leq \operatorname{d}_{\text{str}}(D_{\text{dup}}) \). Finally, the bound \( \operatorname{d}_{\text{str}}(D_{\text{dup}}) \leq \ell_{\text{poly}}(\mathcal{K}_1) + \ell_{\text{poly}}(\mathcal{K}_2) \), follows from applying Phase 1 of FRA-Poly to \( D_{\text{dup}} \).

\( \square \)
We now consider the particular case where $K$ is the doubly nonnegative cone $D^n = S^+_{d} \cap N^n$, where $N^n$ is the cone of $n \times n$ symmetric matrices with nonnegative entries. This cone is important because it can be used as a relatively tractable relaxation for the cone of completely positive matrices, see [26, 8, 2]. The following corollary follows immediately from Theorem 24.

**Corollary 25.** When $K = D^n$, we have $d(D) \leq n$ and $d_{str}(D) \leq n - 1$.

**Proof.** It follows from Theorem 24 by recalling that $\ell_{poly}(S^+_n) = n - 1$ and $\ell_{poly}(N^n) = 0$. 

We will compare the bound in Corollary 25 with the one predicted by the classical FRA. To do that, we need to compute $\ell_{D^n}$.

**Proposition 26.** The longest chain of nonempty faces in $D^n$ has length $\frac{n(n+1)}{2} + 1$, which is the maximum possible for a cone contained in $S^n$.

**Proof.** The assertion about the maximality follows from the fact that if we have two faces such that $F \subseteq F'$, then we must have $\dim(F) < \dim(F')$. Since $S^n$ has dimension $\frac{n(n+1)}{2}$, we cannot have a strictly ascending chain containing more than $\frac{n(n+1)}{2} + 1$ nonempty faces.

Let $G$ be any set of tuples $(i, j)$ with $i, j \in \{1, \ldots, n\}$ and let $N^n(G)$ be the face of $N^n$ which corresponds to the matrices $x$ such that the only entries $x_{i,j}$ that are allowed to be nonzero are the ones for which either $(i, j) \in G$ or $(j, i) \in G$. We will first define two chains of faces of $N^n$. First, let $G_0 = \emptyset$ and define $G_i = G_{i-1} \cup \{(i, i)\}$ for $i \in \{1, \ldots, n\}$. We now consider the following construction written in pseudocode.

```
\begin{align*}
\text{k} &\leftarrow 1, G_0 \leftarrow G_n \\
\text{For} \ i &\leftarrow 1, i \leq n \ \text{do} \\
&\quad \text{For} \ j \leftarrow 1, j < i \ \text{do} \\
&\quad\quad H_k \leftarrow H_{k-1} \cup \{i, j\} \\
&\quad\quad k \leftarrow k + 1 \\
&\quad\quad j \leftarrow j + 1. \\
&\quad i \leftarrow i + 1.
\end{align*}
```

The idea is to add one non-diagonal entry per iteration, so that $N^n(H_k) \subseteq N^n(H_{k+1})$. First $(2, 1)$ will be added, then $(3, 1), (3, 2)$ and so on. We have

$$S^+_n \cap N^n(G_0) \subseteq \cdots \subseteq S^+_n \cap N^n(G_n) \subseteq S^+_n \cap N^n(H_1) \subseteq \cdots \subseteq S^+_n \cap N^n(H_{n(n+1)/2})$$

and all inclusions are indeed strict. The first $n$ inclusions are strict because $S^+_n \cap N^n(G_i) = N^n(G_i)$ and it is clear that $N^n(G_i) \subseteq N^n(G_{i+1})$. Now, let $I_n$ denote the $n \times n$ identity matrix. If $k > 0$ and $x \in \text{ri} N^n(H_k)$ then $x_{i,j} > 0$ for some $(i, j)$ entry such that neither $(i, j)$ nor $(j, i)$ belong to $H_{k-1}$. For $\alpha > 0$ sufficiently large, we have $x + \alpha I_n \in S^+_n \cap N^n(H_k)$ and $x + \alpha I_n \notin S^+_n \cap N^n(H_{k-1})$. This shows the remainder of the containments and concludes the proof, since the chain has length $\frac{n(n+1)}{2} + 1$. 

For feasible problems, the classical FRA analysis gives either the bound $\ell_{D^n} = 1 = \frac{n(n+1)}{2}$ or, using Theorem 24, the bound $\ell_{D^n} + 1 + \ell_{N^n} = 1 = n + \frac{n(n+1)}{2}$. Both bounds are quadratic in $n$ in opposition to the linear bound obtained in Corollary 25.

**References**


Then the following holds:

**Theorem 27** (Rockafellar) said to be a polyhedral function. We need to recall the following theorem on infimal convolution.

Let $\theta$ be a convex function. We denote the domain of $\theta$ by $\text{dom} \theta$. If $\text{dom} \theta$ is not empty and $\theta$ is never $-\infty$, then $\theta$ is said to be proper. Its conjugate will be denoted by $\theta^*$ and it satisfies $\theta^*(s) = \sup_x \langle x, s \rangle - \theta(x)$. If the epigraph of $\theta$ is a polyhedral set, then $\theta$ is said to be a polyhedral function. We need to recall the following theorem on infimal convolution.

**Theorem 27** (Rockafellar). Let $f_1, \ldots, f_m$ be proper convex functions and let $f_{k+1}, \ldots, f_m$ be polyhedral functions. Suppose also that

$$\text{ri} (\text{dom} f_1) \cap \ldots \cap \text{ri} (\text{dom} f_k) \cap \text{dom} f_{k+1} \cap \ldots \cap \text{dom} f_m \neq \emptyset.$$  

Then the following holds:

$$(f_1 + \ldots + f_m)^*(s) = \inf \{f_1^*(s_1) + \ldots + f_m^*(s_m) \mid s_1 + \ldots + s_m = s\},$$

where for each $s$ the infimum is attained whenever it is finite. In other words, under the assumptions of the theorem, the conjugate of the sum is equal to the infimal convolution of the conjugates.

**Proof.** See Theorem 20.1 of Rockafellar [22].

**Proposition 28.** Let $K = K^1 \times K^2$, where $K^1 \subseteq \mathbb{R}^{n_1}, K^2 \subseteq \mathbb{R}^{n_2}$ are closed convex cones such that $K^2$ is polyhedral.

i) If $\theta_P$ is finite and $\square P$ satisfies the PPS condition, then $\theta_P = \theta_D$ and the dual optimal value is attained.

ii) If $\theta_D$ is finite and $\square D$ satisfies the PPS condition, then $\theta_P = \theta_D$ and the primal optimal value is attained.

**Proof.** We will prove i) first. Let $f_1$ be such that $f_1(x) = \langle c, x \rangle$ if $Ax = b$ and $+\infty$ otherwise. Let $f_2$ be the indicator function of $\mathbb{R}^{n_1} \times (K^2)^*$ and $f_3$ be the indicator function of $(K^1)^* \times \mathbb{R}^{n_2}$. Since there is a primal feasible solution $x = (x_1, x_2)$ such that $x_1 \in \text{ri} (K^1)^*$, we have that $\text{dom} f_1 \cap \text{dom} f_2 \cap \text{ri} (\text{dom} f_3)$ is nonempty. In addition, $f_1$ and $f_2$ are polyhedral functions. Let us now observe that:

$$f_1^*(s) = \begin{cases} 
\langle b, y \rangle & \text{if there is } y \text{ with } s - c = A^T y \\
+\infty & \text{otherwise}
\end{cases}$$
Note that, due to feasibility, for fixed $s$, $\langle b, y \rangle$ does not depend on the choice of $y$, as long as $c + A^T y = s$.

This is because since there is $x$ such that $Ax = b$, we have $\langle b, y \rangle = \langle x, s - c \rangle$. The conjugate $f_2^*$ is the indicator function of $-\{0\} \times \mathcal{K}^2$ and $f_3^*$ is the indicator function of $-\mathcal{K}^1 \times \{0\}$. Applying Theorem 27 with $s = 0$, we have:

$$(f_1 + f_2 + f_3)^*(0) = \inf \left\{ \langle b, y \rangle \mid c + A^T y = s_1, s_1 - (0, s_2) - (s_3, 0) = 0, s_2 \in \mathcal{K}^2, s_3 \in \mathcal{K}^1 \right\}$$

$$= \inf \left\{ \langle b, y \rangle \mid c + A^T y = s_1, s_1 \in \mathcal{K}^1 \times \mathcal{K}^2 \right\}$$

$$= -\sup \left\{ \langle b, y \rangle \mid c - A^T y = s_1, s_1 \in \mathcal{K}^1 \times \mathcal{K}^2 \right\},$$

where the sup in the last equation is attained. So, there is some dual feasible $y$ such that $(f_1 + f_2 + f_3)^*(0) = \langle b, y \rangle$. However, using the definition of conjugate, we also have:

$$(f_1 + f_2 + f_3)^*(0) = -\inf \{ \langle c, x \rangle \mid Ax = b, x \in \mathcal{K}^1 \times \mathcal{K}^2 \} = -\theta_P.$$ 

It follows that $\theta_P = \theta_D$ and the dual is attained at $y$. To prove $ii$), let $g_1 = f_1^*$, and let $g_2$ and $g_3$ be the indicator functions of $\mathbb{R}^{n_1} \times \mathcal{K}^2$ and $\mathcal{K}^1 \times \mathbb{R}^{n_2}$, respectively. Again, it is enough to compute $(g_1 + g_2 + g_3)^*(0)$ using both the definition of conjugate function and using Theorem 27. \qed